

Are they answering what we think we're asking?

The importance of pre-testing survey questions using cognitive interviews

Cathryn McCormack
University of New England
Contact: cathryn.mccormack@une.edu.au

Abstract

Evaluation often involves conducting surveys. To create a survey, a typical process consists of: drafting questions using our experience and understanding of basic principles of good question writing to guide us, and obtaining feedback from colleagues. Sometimes a focus group might be included, or conducting a small scale trial. However this process misses the vital pre-testing stage of cognitive interviews. Cognitive interviews are widely used in the development of surveys for health and social sciences, but appear to be underutilised to inform the development of instruments in other fields.

In the process of reviewing its unit and teaching evaluation instruments, the University of New England (UNE) investigated both existing and proposed questions with cognitive interviews using a retrospective think aloud protocol. Over 120 interviews were conducted with students. Each student was asked three or four questions from the full set, resulting in about 20 responses for each question. Interviews were recorded, transcribed, and analysed for variation in understanding. The cognitive interviews identified problems with conceptualisation and phrasing in the question wording, as well as problems with variation of interpretation of the items on the response scale. These problems were not at all clear to the panel developing the survey prior to the interviews being conducted.

From the results of the interviews, the conclusion is made that cognitive interviews are an essential step in the development of survey questions. They reveal problems with the questions previously unnoticed by the survey developers, however expert they may be. In this particular case the retrospective think-aloud protocol was highly worthwhile. Cognitive interviews are considered so valuable by leading authors in the field, that they suggest when a large cognitive interview trial is impractical, it is worth conducting a small number of informal cognitive interviews.

The Survey Development Process

A survey question is a kind of instruction. Respondents are, in effect, instructed to carry out a task – to do whatever mental work is necessary to provide an answer. By this view, they can get into trouble in several ways. For example, respondents can misunderstand what they have been asked to do and, as a result, carry out the wrong task. Or they can understand the instruction just fine but find themselves unable to carry out the task they have been assigned. Or they can successfully carry out the assigned task but find themselves scale survey development process. When unable to fit their answer into the options provided.

(Conrad, Blair, & Tracy, 1999, p.1)

Surveys are one of the tools of the trade of evaluators. Their purpose is to collect information from a large number of people using a structured format. In developing a survey instrument, important phases such as clarifying the survey concept, writing well constructed questions, and field testing the instrument may be employed. However, if respondents are unable to carry out the instructions we ask of them, the information they collect is likely to be inaccurate. To compound the problem, respondents, in their desire to be helpful, are likely to answer a question they don't understand, or select an inappropriate response rather than not applicable.

Cognitive interviews are considered by many statistical agencies as the most effective way of identifying these types of problems in surveys. They have become commonplace in mainstream data collection agencies such as the US Department of Statistics, the US Department of Health, Statistics Canada, Statistics NZ, and even Australia's own Bureau of Statistics, such that cognitive interviews have been a standard part of survey pre-testing. Furthermore, many have their own cognitive interview laboratory facilities to facilitate the testing process.

What is a Cognitive Interview?

The cognitive interview theoretical model is based on the four-stage response model of thought process (Tourangeau, Rips, & Rasinski, 2000), which involves the respondent

- a) comprehending the item
- b) retrieving relevant information
- c) making a judgement based upon knowledge recalled, and
- d) mapping the answer onto the reporting system.

Any of these areas may be problematic for a respondent, and the purpose of the cognitive interview is to identify in which of those areas there is a problem.

Two basic forms of cognitive interviews are the concurrent think-aloud and retrospective think-aloud. The concurrent think-aloud protocol involves the participant talking through their thinking processes as they answer the question. Dillman & Redline (2004, p302) however, question whether a concurrent think-aloud protocol accurately represents how a participant would complete the survey if left to their own devices. In contrast, the retrospective think-aloud protocol allows the interviewer to obtain the participant's first response, then through the retrospective think-aloud process, allow them to explain how they arrived at their answer. The interviewer can probe more deeply as part of either think-aloud protocol, however the retrospective protocol allows the participant's initial response to be recorded. Cognitive interviews are effective only where participants are able to articulate their thinking processes.

Somewhat confusingly, the most common cognitive interview standard is described by Willis (2005) as a concurrent think-aloud, but by others as the immediate retrospective think-aloud. This technique consists of the respondent first answering the question, then being asked by the interviewer to explain how they decided upon their answer. The explanation process usually involves the interviewer using a series of probes to follow up various issues.

Analysis may be as precise as measuring the time delay between the interviewer asking the question and the respondent answering, or as simple as the interviewer noting where respondents encountered problems with the questions. More typically analysis involves some type of coding of responses either in real time or from audio recording or transcript. While Presser & Blair (1994) found that rigorous coding produces the most reproducible results and DeMaio & Landreth that it “may result in a greater understanding of the questionnaire problems” (2004, p.107), DeMaio & Landreth acknowledge a sample of at least 100 responses is required for their rigorous coding analysis to be valid.

Cognitive interviews - unique data collectors

Cognitive interviews, by their nature, are labour intensive and therefore expensive to conduct. However the wealth of information they provide means that a number should be planned into every survey development process. The particular number depends on the size and importance of the survey.

In the course of developing a student learning experience instrument at UNE, I conducted over 150 cognitive interviews, each lasting between 10 and 15 minutes. Each included only three or four questions from the full survey, resulting in between 7 and 20 responses for each question. In some cases, 7 responses was enough to identify a question had problems. Take for instance, the question:

- The unit covered what the unit description said it would (1)
(Strongly Disagree -> Strongly Agree)

Students were asked to initially answer the question, and all of them confidently provided a rating. They were then asked what they considered the unit description to be. Answers ranged from the unit title, the unit introduction (from the learning materials), the entire unit handbook, through to no idea. Only four could correctly identify the unit description as the 100 word piece of text used in the university handbook, and of those only one could remember what it said. This is such a blatant problem it may have been identified in a focus group, but it would not have been identified based solely on a trial or field test.

Another set of questions tested were about assessment. Three variations were tested:

- The assessment tasks assisted me in achieving the learning outcomes (2)
- The assessment tasks provided me with opportunities to fully demonstrate my knowledge and skills (3)
- The assessment tasks assisted my learning in this unit (4)

When presented to a roomful of higher education evaluators at the Evaluation Forum in November 2007, whom one would consider experts, opinions were approximately equally divided between which question had the best wording. The first variation (2) had similar problems than that encountered with the unit description question above (1).

Student: [Reads question] I'd probably say neutral, because I learnt a lot about some areas ...
Interviewer: Did you actually know what the learning outcomes were?
Student: To be honest, probably not really, no.
[Further discussion]
Interviewer: So ... there weren't clearly defined learning outcomes?
Student: Not really.
Interviewer: I'm surprised you gave it as good a score as you did!
Student: Maybe I'm exaggerating the problem ... I was dissatisfied with the whole unit, it wasn't what I was expecting.

The second option (3) resulted in problems with confusion about the meaning of the words 'fully', and the concept of 'demonstrating skills'. The final option encountered almost no problems at all, and we therefore concluded that of the three this was the best choice. Both these examples illustrate how cognitive interviews identified covert problems as well as overt problems with the questions.

Cognitive interviews may also reveal cases where respondents have trouble mapping their response to the scale provided. One common choice of scale is 5 points ranging from strongly disagree through to strongly agree. Our cognitive interviews however, revealed that respondents found it easier to rate using a very poor through excellent scale. Furthermore, the layout of the scale impacted on the ability of students to map their rating to the scale. When words were read out (telephone interview), or were displayed on the hand out card separated only by commas, significant inconsistencies were found when respondents were asked to map their word response on a 0-10 scale. These problems disappeared when the words were spaced and numbered from 1 to 5.

Where cognitive interviews fit in the survey development cycle

The diagram in Attachment 1 illustrates the survey pre-testing process. The first step is to identify the conceptual information required, for instance, the effectiveness of the government's diabetes health risk advertisements, are people eating enough fruit and vegetables, or in our case, what students think about the quality of their learning experience at the University of New England. Concepts may be clarified using focus groups or discussion with stakeholders. A complex survey may require several rounds, but at some stage sufficient information will be available for questions to be drafted.

Draft questions may then be circulated amongst colleagues or a survey expert for feedback, or further focus groups may be conducted. Information collected at this stage may indicate problems with the survey concept, which again has to be revisited. Again, several rounds of feedback collection may be required before the set of survey questions are ready for the next stage.

Pre-testing may consist of field testing, a trial, or cognitive interviews. In practice, these often take place simultaneously with several rounds of each conducted. The amount of pre-testing conducted should be in line with the size and importance of the survey. In my own field, higher education, field testing and cognitive

interviews are rarely undertaken, and testing may consist solely of a trial, or none may be conducted at all. The final result of this is a set of thoroughly test survey questions.

Where the scope of the survey requires it, further testing on aspects of the survey such as print layout or online functionality would also be included in the pre-testing.

How to conduct cognitive interviews

The purpose of the cognitive interview is to discover any problems with the question. The interviewer's role therefore, is to help the respondent talk their way through the answering process. On occasions, coaxing respondents to talk can be a challenge, and as with teaching, asking the right questions is key. Questions can involve asking the respondent to paraphrase the question, explain their logic in calculating an answer (hours per week of study for instance), asking what things they were thinking of when they chose a particular rating, why they didn't rate higher or lower, or asking about the understanding of key words in the question. For a large project where several interviewers are required, consistency across all the interviews becomes more important and the questions may need to be pre-scripted.

While some statistical agencies have access to a cognitive interview laboratory, most survey writers are less well resourced. In practice, this means interviews can be recorded for later transcription or note making, or notes can be made during or immediately after the interview. My personal experience was that my perception of the main points raised in a series of five x 10 minute interviews proved to be inaccurate when compared with the recordings. My recommendation would be therefore to either record the interview for later playback, or have an assistant make notes while the interview is taking place.

When recruiting for cognitive interviews, participants need to be both representative of the group the survey is designed for, and consist of the widest possible range of members of that group. In our testing we recruited both campus based and distance education students, and students from as many disciplines and unit structures as possible.

Limitations of cognitive interviews

Cognitive interviews are a qualitative research method, and respondents can only be considered as a series of case studies. It is not appropriate to report for instance, that 8 out of 10 respondents encountered a particular problem. In some instances, a single case may be adequate to identify particular problems with a question. The survey developer must use their discretion to consider how widespread a particular problem might be. Within our testing for instance, where a single student identified a question as unanswerable based on the unit organisation, we could use our knowledge of how units are organised at the university and make a judgement as to how frequently this situation would likely occur.

Conclusions

Cognitive interviews are so valuable to pre-testing interview questions, that even when there is little time and funding to conduct them, any number is better than none. For a large scale, high stakes survey, several rounds of up to ten may be required. Where a survey is long, questions can be divided into sub-groups for testing to facilitate shorter interview times.

One final note: gatekeeper questions, that is, questions that guide respondents into different sections of the survey require particular attention in the testing process.

If you are interested in building your knowledge further in this area, I particularly recommend (Willis, 2005). It is a highly practical book covering all aspects of conducting cognitive interviews.

References

- Conrad, F., Blair, J., & Tracy, E. (1999). *Verbal Reports are Data! A Theoretical Approach to Cognitive Interviews*. Paper presented at the Federal Committee on Statistical Methodology, Washington DC: Office of Management and Budget.
- DeMaio, T. J., & Landreth, A. (2004). Do Different Cognitive Interview Techniques Produce Different Results? In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin & E. Singer (Eds.), *Methods for Testing and Evaluating Survey Questionnaires* (pp. 89-108). Hoboken, New Jersey: Wiley.
- Dillman, D. A., & Redline, C. D. (2004). Testing Paper Self-Administered Questionnaires: Cognitive Interview and Field Test Comparisons. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin & E. Singer (Eds.), *Methods for Testing and Evaluating Survey Questionnaires*: Wiley Interscience.
- Presser, S., & Blair, J. (1994). Survey Pretesting: Do Different Methods Produce Different Results? *Sociological Methodology*, 24, 73-104.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.
- Willis, G. B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks: Sage Publications Inc.

Attachment 1

Survey Development Diagram

